In [ ]:
```
import pandas as pd
```

In [ ]:
```
import pandas as pd
import os
import seaborn as sns
sns.set(color_codes=True)
import matplotlib.pyplot as plt
import plotly.graph_objects as go
import glob
import numpy as np
import warnings
import requests
import plotly.express as px
```

In [47]:
```
pwd
```

Out[47]:
```
'C:\\Users\\Kareem'
```

In [48]:
```
variable = pd.read_csv(r"C:\Users\Kareem\Desktop\AIML ASSGINMENTS\Online food\food.csv
variable.head()
```

Out[48]:

	Age	Gender	Marital Status	Occupation	Monthly Income	Educational Qualifications	Family size	latitude	longitude	Pin code	C
0	20	Female	Single	Student	No Income	Post Graduate	4	12.9766	77.5993	560001	
1	24	Female	Single	Student	Below Rs.10000	Graduate	3	12.9770	77.5773	560009	
2	22	Male	Single	Student	Below Rs.10000	Post Graduate	3	12.9551	77.6593	560017	
3	22	Female	Single	Student	No Income	Graduate	6	12.9473	77.5616	560019	
4	22	Male	Single	Student	Below Rs.10000	Post Graduate	4	12.9850	77.5533	560010	

In [49]:
```
variable['Unnamed: 12'].value_counts()
```

Out[49]:
```
Unnamed: 12
Yes    301
No      87
Name: count, dtype: int64
```

In [50]:
```
variable.drop(columns=['Unnamed: 12'],inplace = True)
variable.head()
```

Out[50]:

	Age	Gender	Marital Status	Occupation	Monthly Income	Educational Qualifications	Family size	latitude	longitude	Pin code	C
0	20	Female	Single	Student	No Income	Post Graduate	4	12.9766	77.5993	560001	
1	24	Female	Single	Student	Below Rs.10000	Graduate	3	12.9770	77.5773	560009	
2	22	Male	Single	Student	Below Rs.10000	Post Graduate	3	12.9551	77.6593	560017	
3	22	Female	Single	Student	No Income	Graduate	6	12.9473	77.5616	560019	
4	22	Male	Single	Student	Below Rs.10000	Post Graduate	4	12.9850	77.5533	560010	

In [51]:
```
variable.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 388 entries, 0 to 387
Data columns (total 12 columns):
 #   Column                       Non-Null Count  Dtype  
---  ------                       --------------  -----  
 0   Age                          388 non-null    int64  
 1   Gender                       388 non-null    object 
 2   Marital Status               388 non-null    object 
 3   Occupation                   388 non-null    object 
 4   Monthly Income               388 non-null    object 
 5   Educational Qualifications   388 non-null    object 
 6   Family size                  388 non-null    int64  
 7   latitude                     388 non-null    float64
 8   longitude                    388 non-null    float64
 9   Pin code                     388 non-null    int64  
 10  Output                       388 non-null    object 
 11  Feedback                     388 non-null    object 
dtypes: float64(2), int64(3), object(7)
memory usage: 36.5+ KB
```

In [52]:
```
variable['Gender'].unique()
```

Out[52]:
```
array(['Female', 'Male'], dtype=object)
```

In [54]:
```
print(variable.head())
```

```
   Age  Gender Marital Status Occupation  Monthly Income  \
0   20       0         Single    Student       No Income
1   24       0         Single    Student  Below Rs.10000
2   22       1         Single    Student  Below Rs.10000
3   22       0         Single    Student       No Income
4   22       1         Single    Student  Below Rs.10000

  Educational Qualifications  Family size  latitude  longitude  Pin code  \
0              Post Graduate            4   12.9766    77.5993    560001
1                   Graduate            3   12.9770    77.5773    560009
2              Post Graduate            3   12.9551    77.6593    560017
3                   Graduate            6   12.9473    77.5616    560019
4              Post Graduate            4   12.9850    77.5533    560010

  Output   Feedback
0    Yes   Positive
1    Yes   Positive
2    Yes  Negative
3    Yes   Positive
4    Yes   Positive
```

In [55]:
```
print(variable.columns)
```

```
Index(['Age', 'Gender', 'Marital Status', 'Occupation', 'Monthly Income',
       'Educational Qualifications', 'Family size', 'latitude', 'longitude',
       'Pin code', 'Output', 'Feedback'],
      dtype='object')
```

In [56]:
```
print(variable['Gender'].unique())
```

```
[0 1]
```

In [57]:
```
print(variable['Gender'].head())
```

```
0    0
1    0
2    1
3    0
4    1
Name: Gender, dtype: int64
```

In [58]:
```
variable['Occupation'].unique()
```

Out[58]:
```
array(['Student', 'Employee', 'Self Employeed', 'House wife'],
      dtype=object)
```

In [59]:
```
variable['Marital Status'].unique()
```

Out[59]:
```
array(['Single', 'Married', 'Prefer not to say'], dtype=object)
```

In [94]:
```
variable['Marital Status'] = variable['Marital Status'].replace({'Single': 1, 'Married
```

In [61]:
```
variable['Occupation'].unique()
```

Out[61]:
```
array(['Student', 'Employee', 'Self Employeed', 'House wife'],
      dtype=object)
```

In [62]:
```
variable['Monthly Income'].unique()
```

Out[62]:
```
array(['No Income', 'Below Rs.10000', 'More than 50000', '10001 to 25000',
       '25001 to 50000'], dtype=object)
```

In [131…
```
variable['Monthly Income'] = variable['Monthly Income'].replace({
    'No Income': 1,
    'Below Rs.10000': 2,
    'More than 50000': 0,
    '10001 to 25000': 4,
    '25001 to 50000': 3
})
```

In [ ]:
```
print(variable.columns)
variable.drop(columns=['Pin code'], inplace=True)
```

In [65]:
```
variable.head()
```

Out[65]:

	Age	Gender	Marital Status	Occupation	Monthly Income	Educational Qualifications	Family size	latitude	longitude	Output	F
0	20	0	1	Student	1	Post Graduate	4	12.9766	77.5993	Yes	
1	24	0	1	Student	2	Graduate	3	12.9770	77.5773	Yes	
2	22	1	1	Student	2	Post Graduate	3	12.9551	77.6593	Yes	
3	22	0	1	Student	1	Graduate	6	12.9473	77.5616	Yes	
4	22	1	1	Student	2	Post Graduate	4	12.9850	77.5533	Yes	

In [67]:
```
variable['Educational Qualifications'].unique()
```

Out[67]:
```
array(['Post Graduate', 'Graduate', 'Ph.D', 'Uneducated', 'School'],
      dtype=object)
```

In [96]:
```
variable['Educational Qualifications'] = variable['Educational Qualifications'].replac
    'Post Graduate': 1,
    'Graduate': 2,
    'Ph.D': 0,
    'Uneducated': 4,
    'School': 3
})
```

In [69]:
```
variable['Output'].value_counts()
```

Out[69]:
```
Output
Yes    301
No      87
Name: count, dtype: int64
```

In [72]:
```
variable.head()
```

Out[72]:

	Age	Gender	Marital Status	Occupation	Monthly Income	Educational Qualifications	Family size	latitude	longitude	Output	F
0	20	0	1	Student	1	1	4	12.9766	77.5993	1	
1	24	0	1	Student	2	2	3	12.9770	77.5773	1	
2	22	1	1	Student	2	1	3	12.9551	77.6593	1	
3	22	0	1	Student	1	2	6	12.9473	77.5616	1	
4	22	1	1	Student	2	1	4	12.9850	77.5533	1	

In [73]:
```
variable['Feedback'].unique()
```

Out[73]:
```
array(['Positive', 'Negative '], dtype=object)
```

In [ ]:
```
variable['Feedback'] = variable['Feedback'].replace({'Negative': 0}, inplace=True)
print(variable.columns)
```

In [ ]:
```
variable.drop(columns= ['Feedback'],inplace  = True)
```

In [76]:
```
variable.isna().sum() # no null values.
```

Out[76]:
```
Age                            0
Gender                         0
Marital Status                 0
Occupation                     0
Monthly Income                 0
Educational Qualifications     0
Family size                    0
latitude                       0
longitude                      0
Output                         0
dtype: int64
```

In [ ]:
```
# who order more food male or female?
```

In [77]:
```
variable.head()
```

Out[77]:

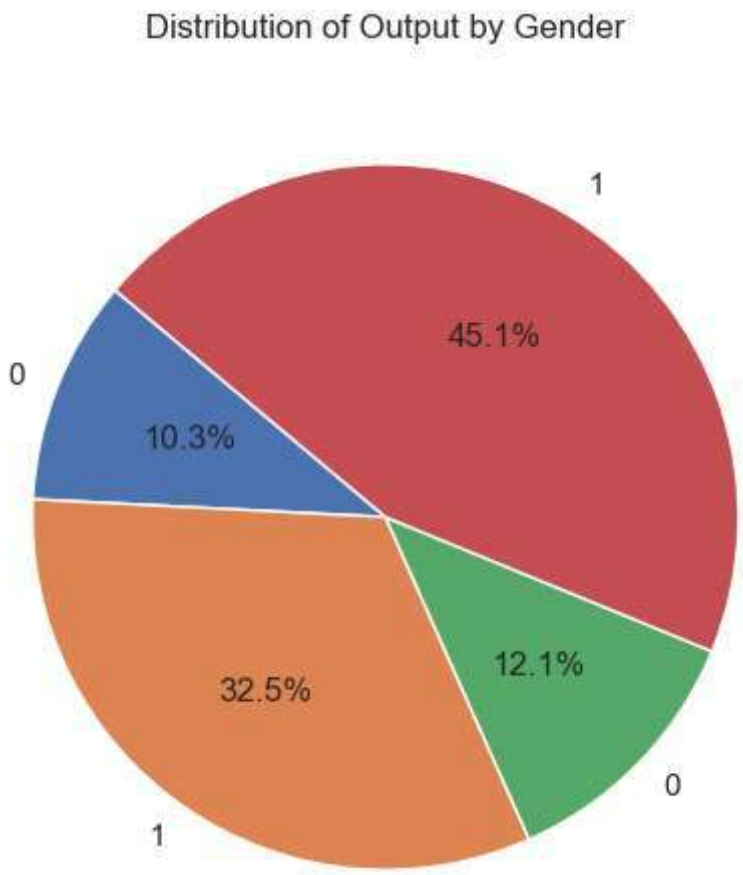
	Age	Gender	Marital Status	Occupation	Monthly Income	Educational Qualifications	Family size	latitude	longitude	Output
0	20	0	1	Student	1	1	4	12.9766	77.5993	1
1	24	0	1	Student	2	2	3	12.9770	77.5773	1
2	22	1	1	Student	2	1	3	12.9551	77.6593	1
3	22	0	1	Student	1	2	6	12.9473	77.5616	1
4	22	1	1	Student	2	1	4	12.9850	77.5533	1

In [80]:
```
male = variable['Gender'].value_counts()[0]
female = variable['Gender'].value_counts()[1]
```

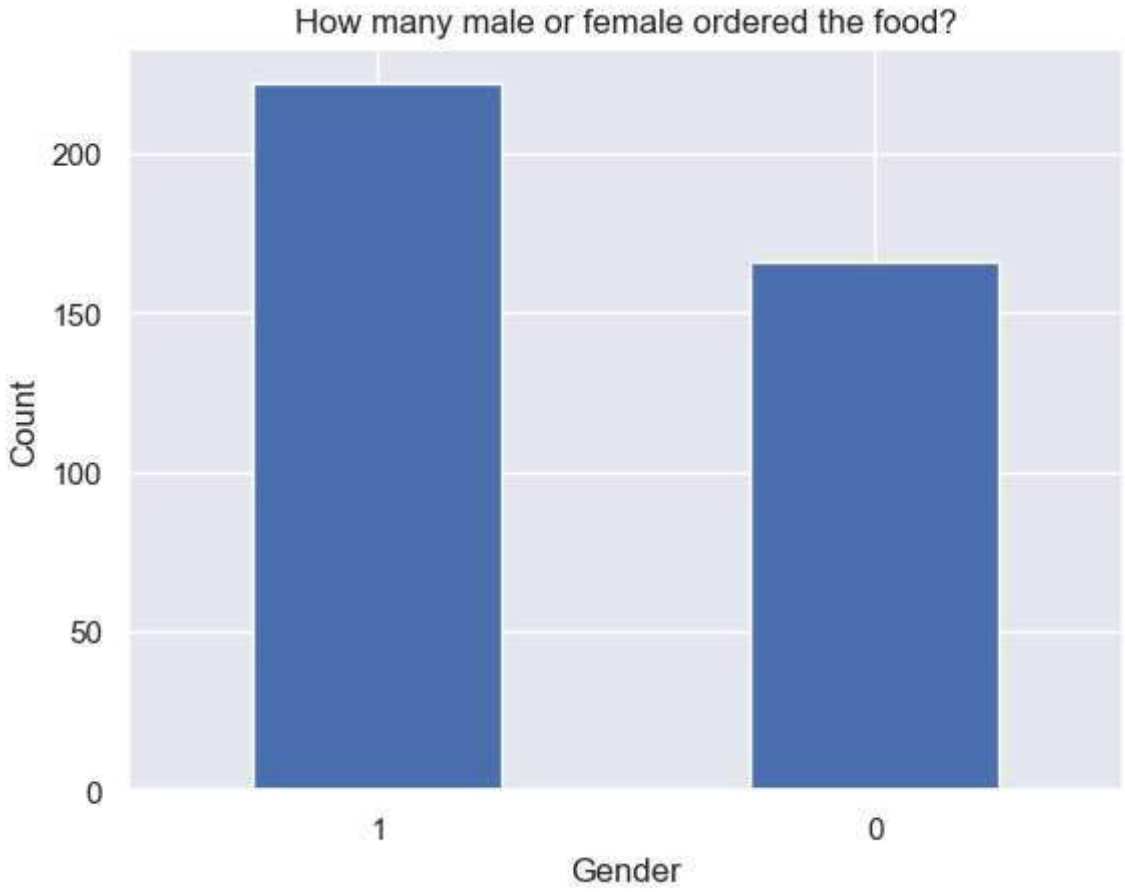
In [81]:
```
male,female
```

Out[81]:
```
(166, 222)
```

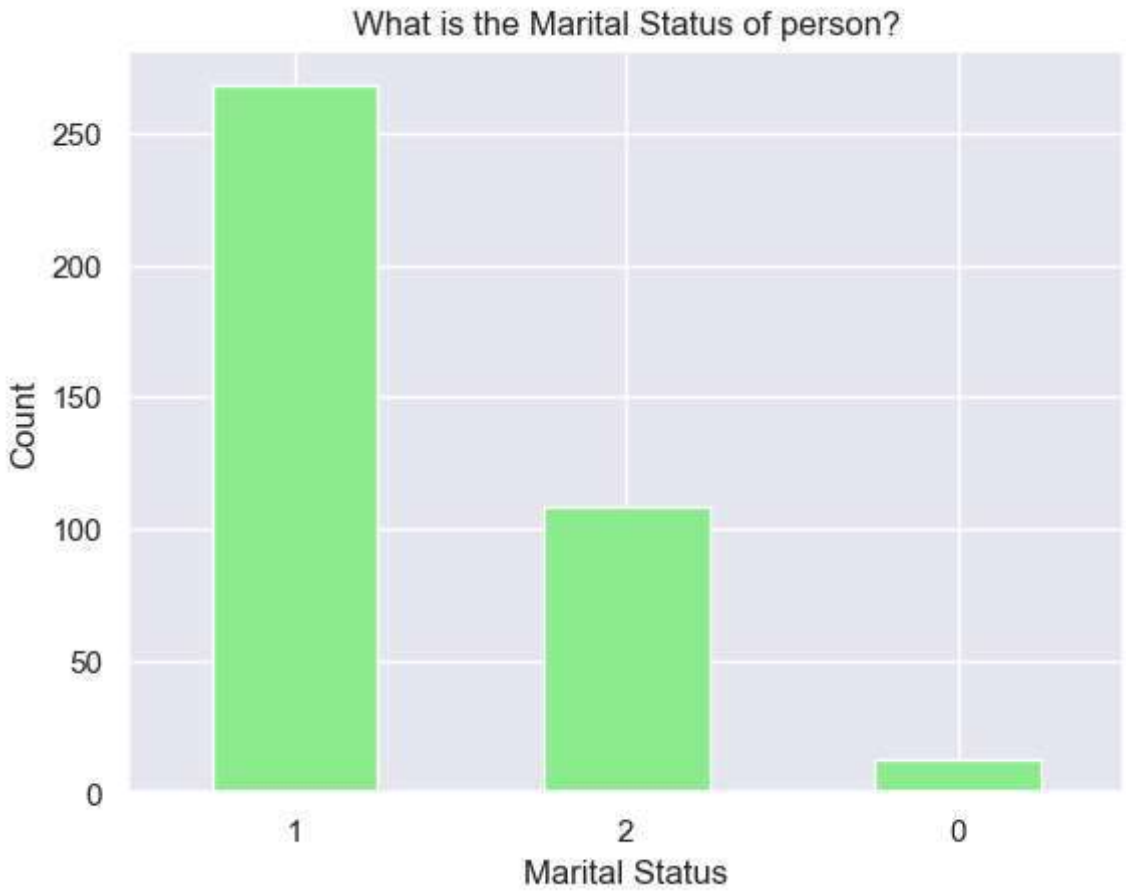
In [84]:
```
gender_output_counts = variable.groupby(['Gender', 'Output']).size().reset_index(name=

plt.figure(figsize=(5, 6))
plt.pie(gender_output_counts['count'], labels=gender_output_counts['Output'], autopct=
plt.title('Distribution of Output by Gender')
plt.axis('equal')
plt.show()
```



In [85]:
```
gender_counts = variable['Gender'].value_counts()

gender_counts.plot(kind='bar', rot=0)
plt.xlabel('Gender')
plt.ylabel('Count')
plt.title('How many male or female ordered the food?')
plt.show()
```



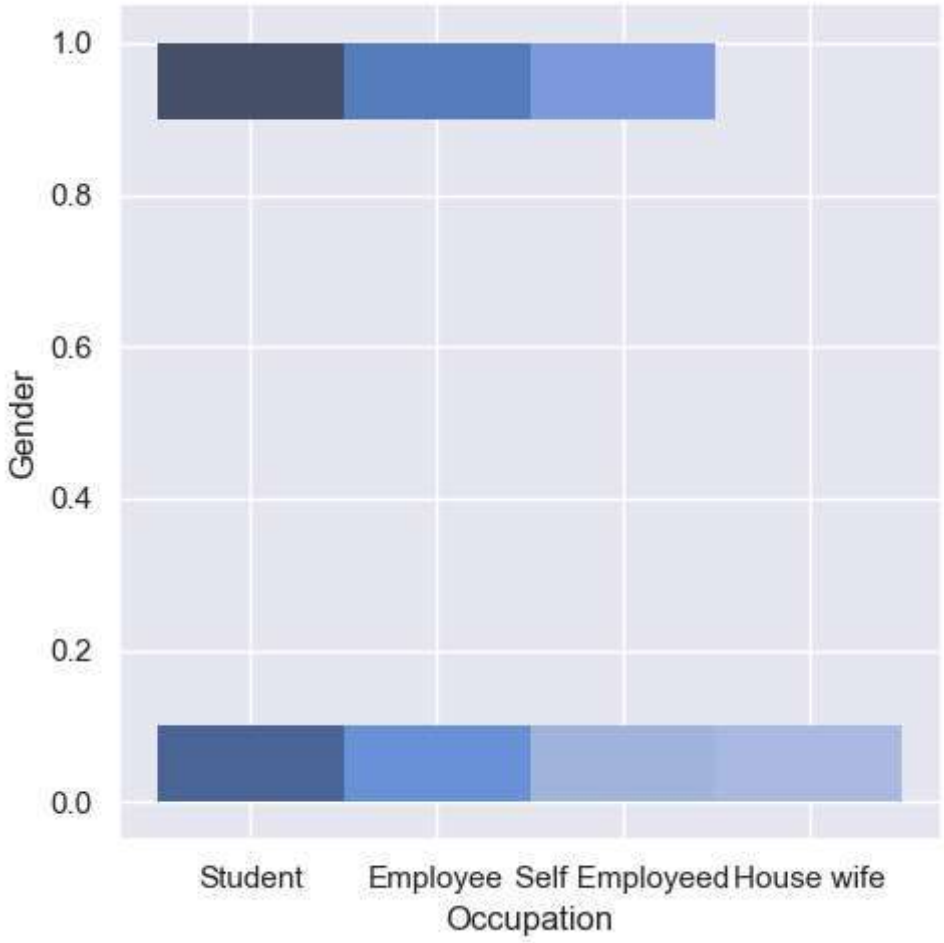
In [86]:
```
MaritalStatus = variable['Marital Status'].value_counts()
MaritalStatus.plot(kind='bar', rot=0, color='lightgreen')
plt.xlabel('Marital Status')
plt.ylabel('Count')
plt.title('What is the Marital Status of person?')
plt.show()
```



In [91]:
```
import seaborn as sns

sns.displot(data=variable, x='Occupation', y='Gender')
```

```
C:\Users\Kareem\anaconda3\Lib\site-packages\seaborn\axisgrid.py:118: UserWarning: The
figure layout has changed to tight
  self._figure.tight_layout(*args, **kwargs)
```

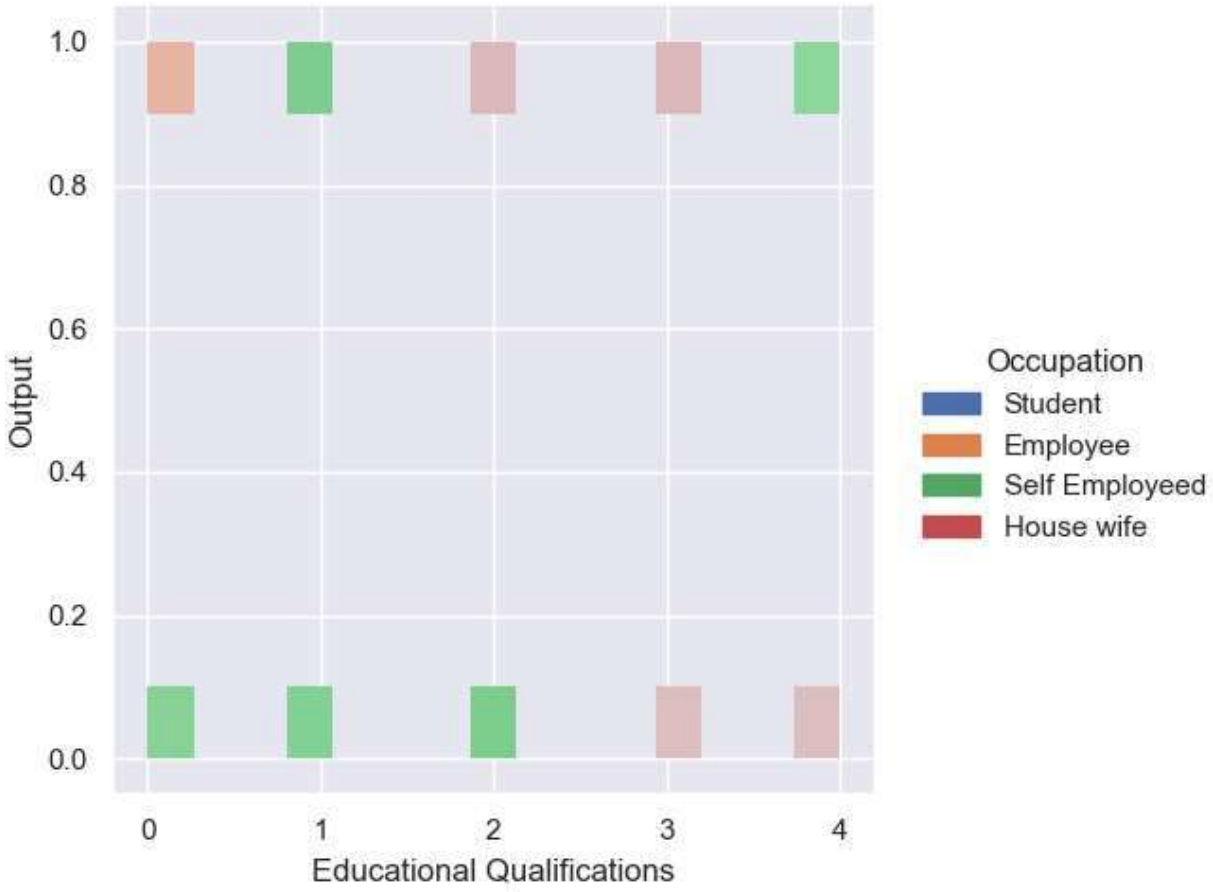
Out[91]:
```
<seaborn.axisgrid.FacetGrid at 0x208c14652d0>
```



In [97]:
```
sns.displot(x='Educational Qualifications', y='Output', data=variable, hue = 'Occupati
```

```
C:\Users\Kareem\anaconda3\Lib\site-packages\seaborn\axisgrid.py:118: UserWarning: The
figure layout has changed to tight
  self._figure.tight_layout(*args, **kwargs)
```

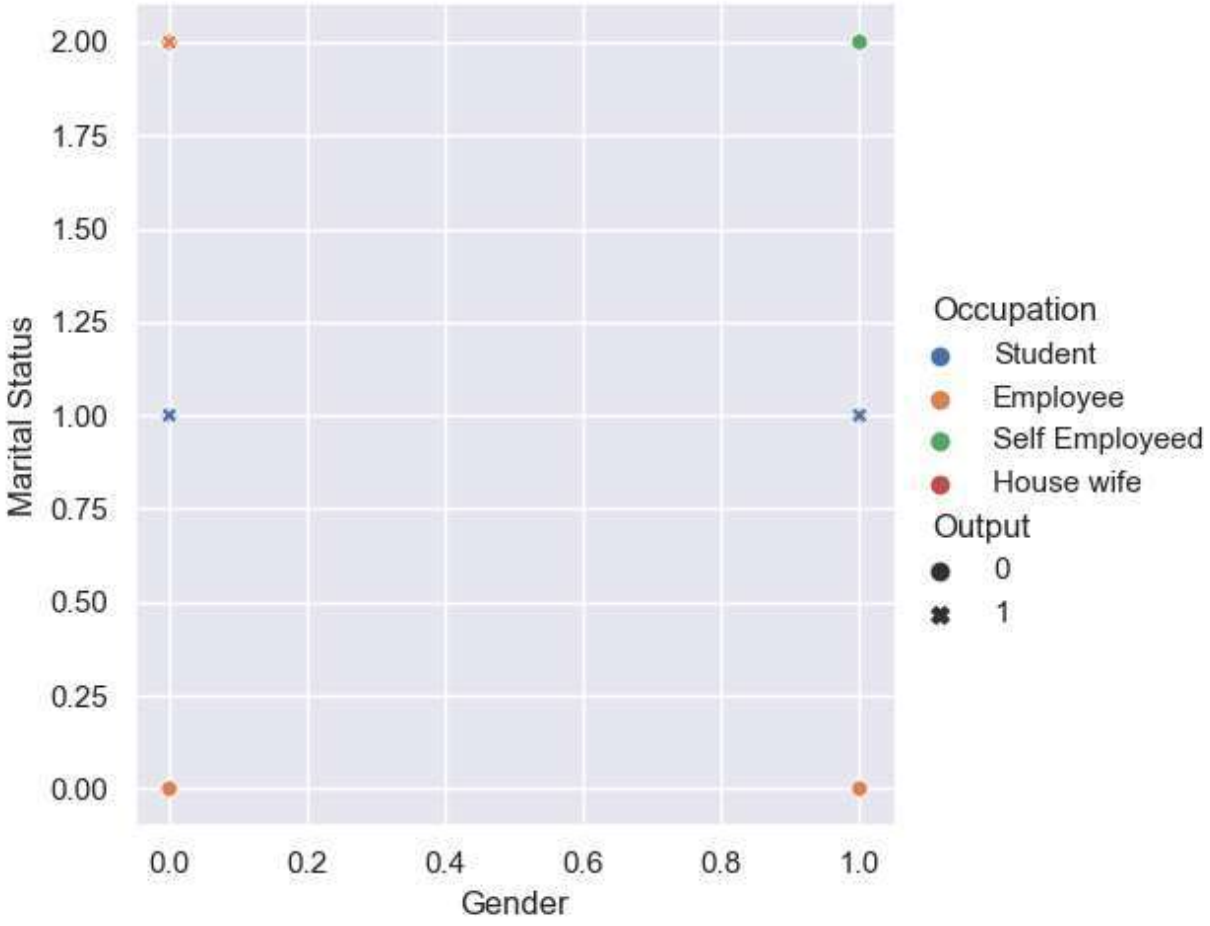
Out[97]:
```
<seaborn.axisgrid.FacetGrid at 0x208c22ae890>
```



In [98]:
```
sns.relplot(x='Gender', y='Marital Status', data=variable, hue = 'Occupation',style= '
```

```
C:\Users\Kareem\anaconda3\Lib\site-packages\seaborn\axisgrid.py:118: UserWarning: The
figure layout has changed to tight
  self._figure.tight_layout(*args, **kwargs)
```

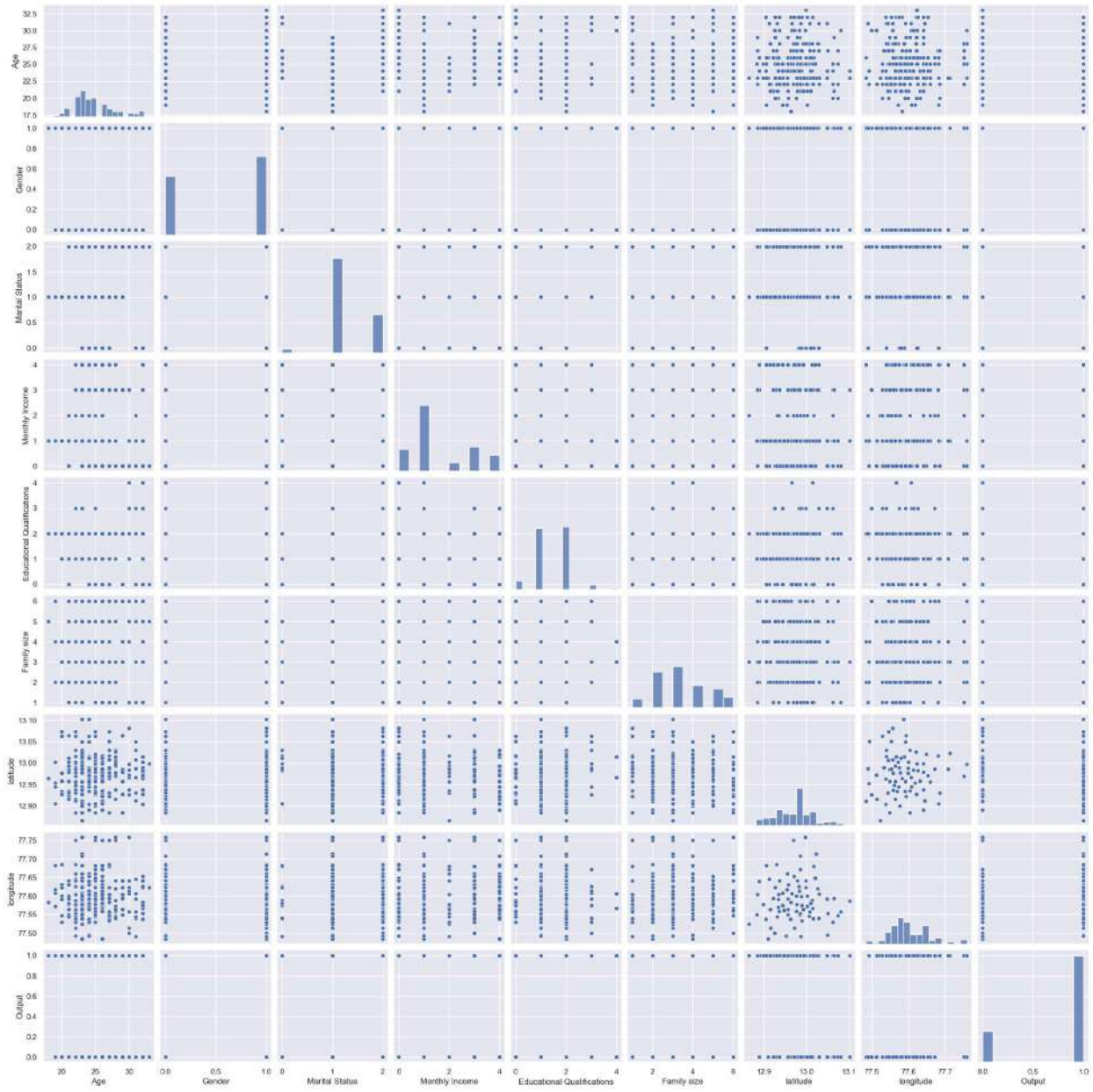
Out[98]:
```
<seaborn.axisgrid.FacetGrid at 0x208b783a150>
```



In [110…
```
import seaborn as sns
sns.pairplot(variable)
```

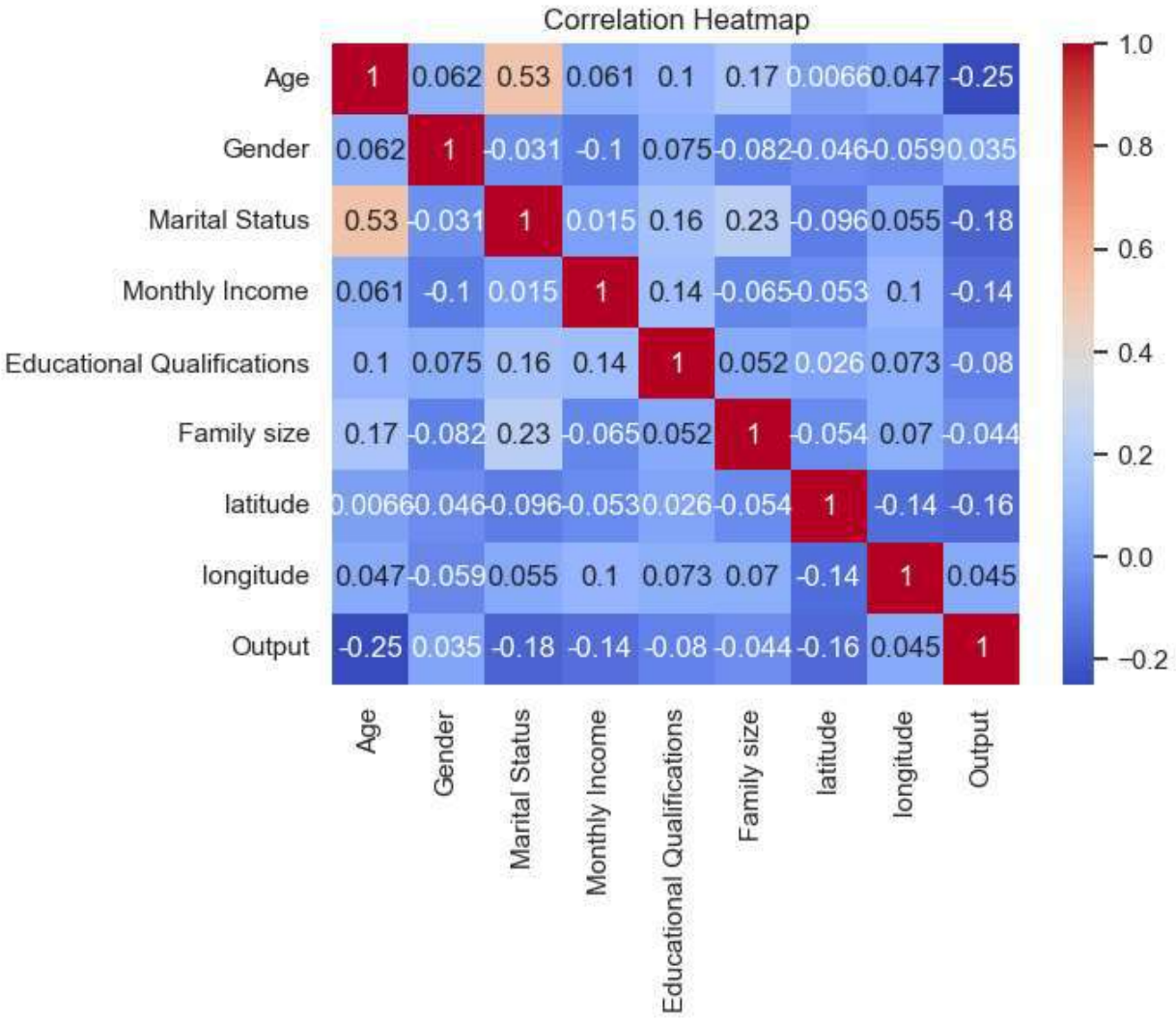
```
C:\Users\Kareem\anaconda3\Lib\site-packages\seaborn\axisgrid.py:118: UserWarning: The
figure layout has changed to tight
  self._figure.tight_layout(*args, **kwargs)
```

Out[110]:
```
<seaborn.axisgrid.PairGrid at 0x208db668fd0>
```



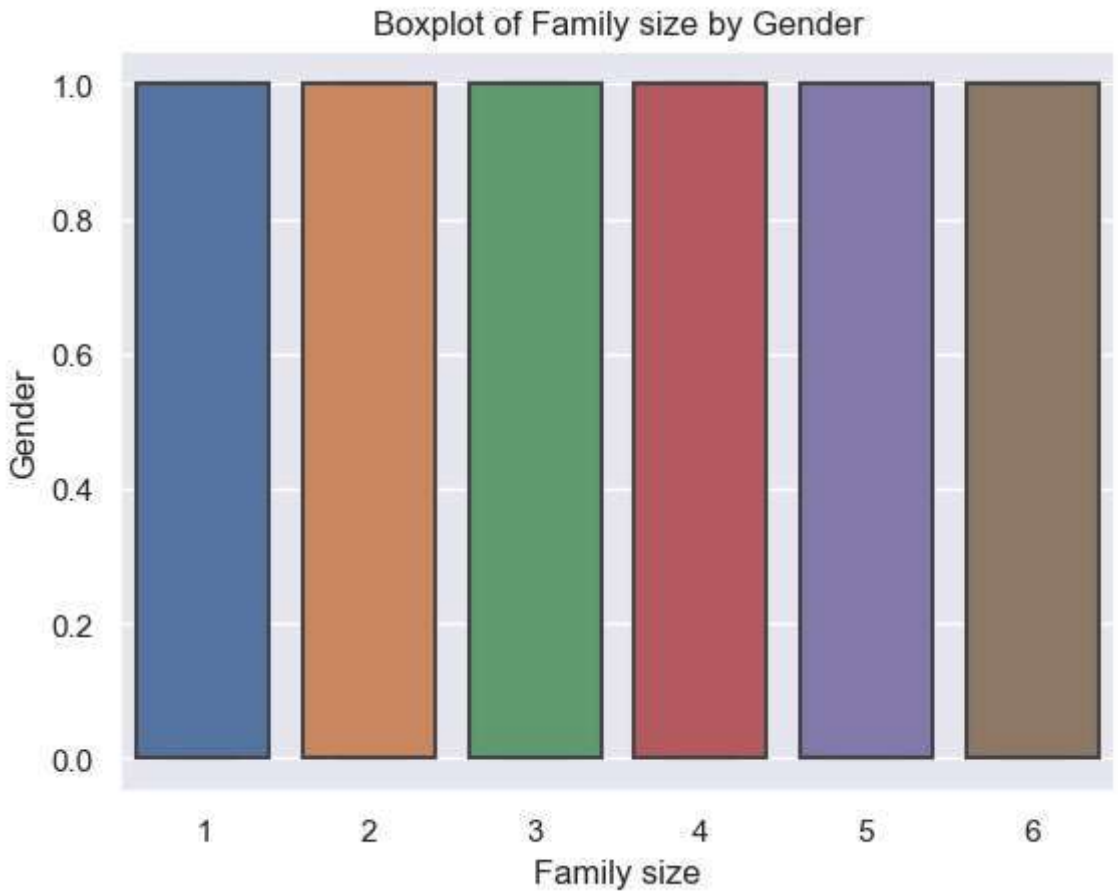
In [114…

```
import seaborn as sns
import matplotlib.pyplot as plt

numeric_variable = variable.select_dtypes(include='number')

correlation_matrix = numeric_variable.corr()

sns.heatmap(correlation_matrix, annot=True, cmap='coolwarm')

plt.title('Correlation Heatmap')

plt.show()
```



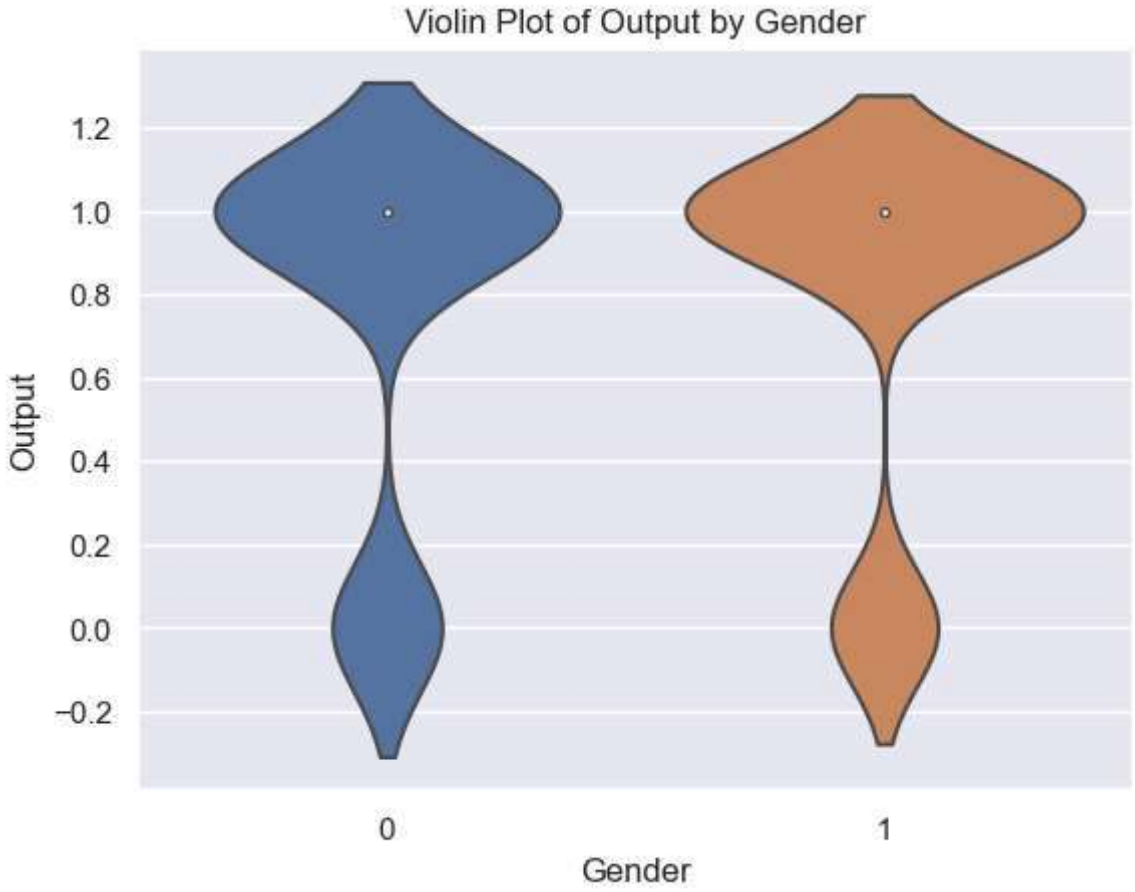
In [116…

```
import seaborn as sns

sns.boxplot(x='Family size', y='Gender', data=variable)

plt.title('Boxplot of Family size by Gender')

plt.show()
```



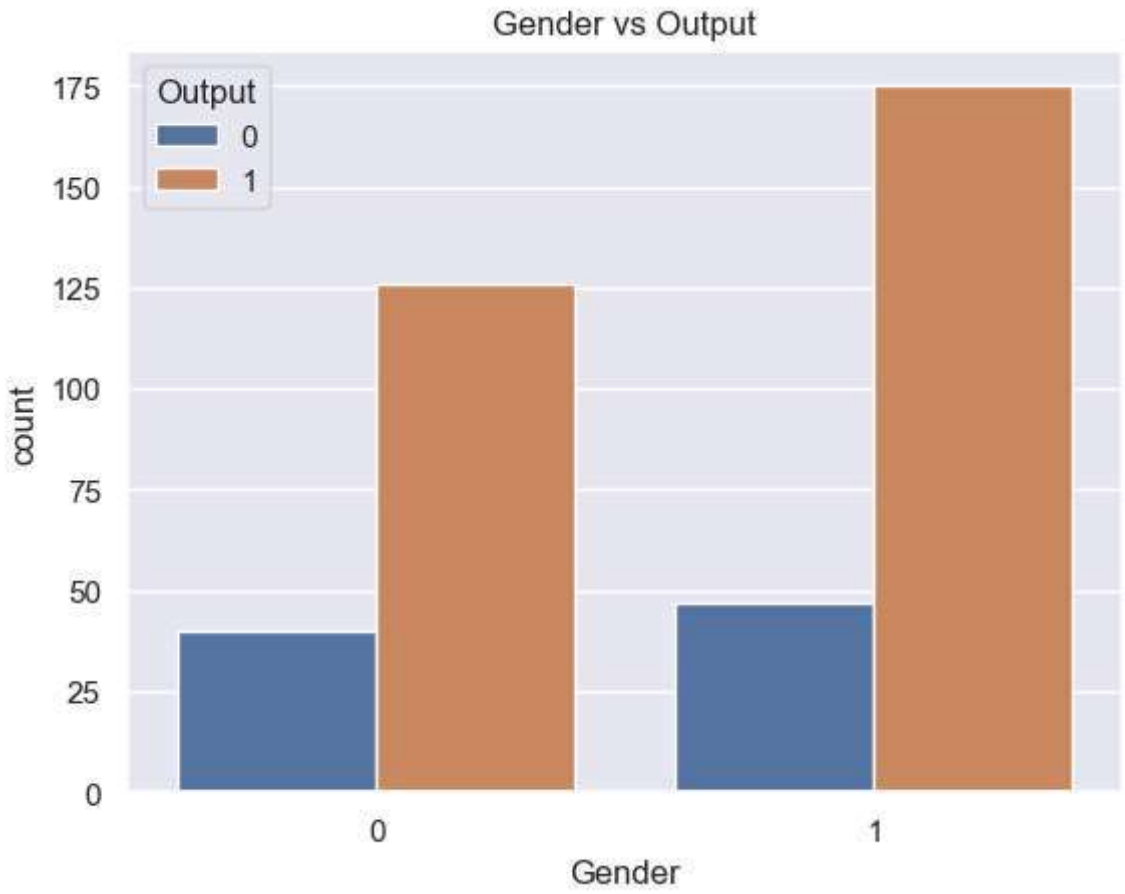
In [117…

```
sns.violinplot(x='Gender', y='Output', data=variable)

plt.title('Violin Plot of Output by Gender')

plt.show()
```



In [118…

```
sns.countplot(x='Gender', hue='Output', data=variable)

plt.title('Gender vs Output')

plt.show()
```


Gender vs Output
Output
0
1
175
150
125
100
75
50
25
0
count
0
1
Gender